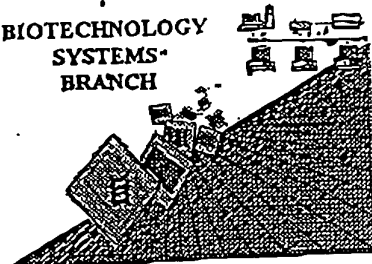Re-run

# RAW SEQUENCE LISTING
# ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/719261_

Source: _PCT_

Date Processed by STIC: _08 30 2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

## Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

**ERROR DETECTED**    **SUGGESTED CORRECTION**    SERIAL NUMBER: _09/719261_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 ____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 ____ Misaligned Amino
Numbering

The numbering under each 5ᵗʰ amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 ____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ____ Variable Length.

Sequence(s) 5 contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 ____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 ____ Skipped Sequences
(OLD RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)        SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 ____ Skipped Sequences
(NEW RULES)

Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 ____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 ____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 ____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 ____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 ____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

The type of errors shown may exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

RAW SEQUENCE LISTING        DATE: 08/30/2001
PATENT APPLICATION: US/09/719,261    TIME: 09:46:14

Input Set : A:\Mewburn-Matassa ('261) Sequence Listing.txt
Output Set: N:\CRF3\08302001\I719261.raw

```
  4 <110> APPLICANT: Istituto Di Ricerche Di Biologia Molecolare P. Angeletti SpA
  5        Matassa, Victor
  6        Narjes, Frank
  7        Koehler, Konrad
  8        Ontoria, Jesus
  9        Poma, Marco
 11 <120> TITLE OF INVENTION: Peptide inhibitors of hepatitis C virus NS3 protease
 13 <130> FILE REFERENCE: KMN/FP5780044
 15 <140> CURRENT APPLICATION NUMBER: 09/719,261
 16 <141> CURRENT FILING DATE: 2001-07-23
 18 <150> PRIOR APPLICATION NUMBER: PCT/GB99/01824
 19 <151> PRIOR FILING DATE: 1999-06-09
 21 <150> PRIOR APPLICATION NUMBER: GB 9812523.0
 22 <151> PRIOR FILING DATE: 1998-06-10
 24 <160> NUMBER OF SEQ ID NOS: 13
 26 <170> SOFTWARE: PatentIn Ver. 2.1
 29 <210> SEQ ID NO: 1
 30 <211> LENGTH: 4
 31 <212> TYPE: PRT
 32 <213> ORGANISM: Artificial Sequence
 34 <220> FEATURE:
 35 <221> NAME/KEY: SITE
 36 <222> LOCATION: (1)
 37 <223> OTHER INFORMATION: Xaa is diphenylalanine
 39 <220> FEATURE:
 40 <221> NAME/KEY: SITE
 41 <222> LOCATION: (3)
 42 <223> OTHER INFORMATION: Xaa is cyclohexylalanine
 44 <220> FEATURE:
 45 <221> NAME/KEY: SITE
 46 <222> LOCATION: (4)
 47 <223> OTHER INFORMATION: Xaa is 4,4-difluoro-2-amino butyric acid
 49 <220> FEATURE:
 50 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 51        sequence
 53 <400> SEQUENCE: 1
 54 Xaa Glu Xaa Xaa
 55  1
 59 <210> SEQ ID NO: 2
 60 <211> LENGTH: 6
 61 <212> TYPE: PRT
 62 <213> ORGANISM: Artificial Sequence
 64 <220> FEATURE:
 65 <221> NAME/KEY: SITE
 66 <222> LOCATION: (6)
 67 <223> OTHER INFORMATION: Xaa is 4,4-difluoro-2-amino butyric acid
 69 <220> FEATURE:
```

RAW SEQUENCE LISTING         DATE: 08/30/2001
PATENT APPLICATION: US/09/719,261    TIME: 09:46:14

Input Set : A:\Mewburn-Matassa ('261) Sequence Listing.txt
Output Set: N:\CRF3\08302001\I719261.raw

```
 70 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 71       sequence
 73 <400> SEQUENCE: 2
W--> 74 Asp Glu Met Glu Glu Xaa
 75   1              5
 78 <210> SEQ ID NO: 3
 79 <211> LENGTH: 6
 80 <212> TYPE: PRT
 81 <213> ORGANISM: Artificial Sequence
 83 <220> FEATURE:
 84 <221> NAME/KEY: SITE
 85 <222> LOCATION: (3)
 86 <223> OTHER INFORMATION: Xaa is diphenylalanine
 88 <220> FEATURE:
 89 <221> NAME/KEY: SITE
 90 <222> LOCATION: (5)
 91 <223> OTHER INFORMATION: Xaa is cyclohexylalanine
 93 <220> FEATURE:
 94 <221> NAME/KEY: SITE
 95 <222> LOCATION: (6)
 96 <223> OTHER INFORMATION: Xaa is 4,4-difluoro-2-amino butyric acid
 98 <220> FEATURE:
 99 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
100       sequence
102 <400> SEQUENCE: 3
W--> 103 Asp Glu Xaa Glu Xaa Xaa
104   1              5
107 <210> SEQ ID NO: 4
108 <211> LENGTH: 6
109 <212> TYPE: PRT
110 <213> ORGANISM: Artificial Sequence
112 <220> FEATURE:
113 <221> NAME/KEY: SITE
114 <222> LOCATION: (3)
115 <223> OTHER INFORMATION: Xaa is diphenylalanine
117 <220> FEATURE:
118 <221> NAME/KEY: SITE
119 <222> LOCATION: (5)
120 <223> OTHER INFORMATION: Xaa is cyclohexylalanine
122 <220> FEATURE:
123 <221> NAME/KEY: SITE
124 <222> LOCATION: (6)
125 <223> OTHER INFORMATION: Xaa is a fluorinated hydrocarbon side chain
127 <220> FEATURE:
128 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
129       sequence
131 <400> SEQUENCE: 4
W--> 132 Asp Glu Xaa Glu Xaa Xaa
133   1              5
```

RAW SEQUENCE LISTING
PATENT APPLICATION:  US/09/719,261

DATE: 08/30/2001
TIME: 09:46:14

Input Set : A:\Mewburn-Matassa ('261) Sequence Listing.txt
Output Set: N:\CRF3\08302001\I719261.raw

```
136 <210> SEQ ID NO: 5
137 <211> LENGTH: 6
138 <212> TYPE: PRT
139 <213> ORGANISM: Artificial Sequence
141 <220> FEATURE:
142 <221> NAME/KEY: SITE
143 <222> LOCATION: (6)
144 <223> OTHER INFORMATION: Xaa is a fluorinated hydrocarbon side chain
146 <220> FEATURE:
147 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
148       sequence
150 <400> SEQUENCE: 5
151 Asp Glu Met Glu Glu Xaa
152   1               5
155 <210> SEQ ID NO: 6
156 <211> LENGTH: 5
157 <212> TYPE: PRT
158 <213> ORGANISM: Artificial Sequence
160 <220> FEATURE:
161 <221> NAME/KEY: SITE
162 <222> LOCATION: (1)
163 <223> OTHER INFORMATION: Asp as tertiary butyl ester
165 <220> FEATURE:
166 <221> NAME/KEY: SITE
167 <222> LOCATION: (2, 4, 5)
168 <223> OTHER INFORMATION: Glu as tertiary butyl ester
170 <220> FEATURE:
171 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
172       sequence
174 <400> SEQUENCE: 6
175 Asp Glu Met Glu Glu
176   1               5
179 <210> SEQ ID NO: 7
180 <211> LENGTH: 5
181 <212> TYPE: PRT
182 <213> ORGANISM: Artificial Sequence
184 <220> FEATURE:
185 <221> NAME/KEY: SITE
186 <222> LOCATION: (1)
187 <223> OTHER INFORMATION: Asp as tertiary butyl ester
190 <220> FEATURE:
191 <221> NAME/KEY: SITE
192 <222> LOCATION: (2, 4)
193 <223> OTHER INFORMATION: Glu as tertiary butyl ester
195 <220> FEATURE:
196 <221> NAME/KEY: SITE
197 <222> LOCATION: (3)
198 <223> OTHER INFORMATION: Xaa is diphenylalanine
200 <220> FEATURE:
```

*[handwritten annotations: "Xaa may only represent a single amino acid", "Errored"]*

W--> 151

*[handwritten note]* The type of errors shown may exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/719,261

DATE: 08/30/2001
TIME: 09:46:14

Input Set : A:\Mewburn-Matassa ('261) Sequence Listing.txt
Output Set: N:\CRF3\08302001\I719261.raw

```
201 <221> NAME/KEY: SITE
202 <222> LOCATION: (5)
203 <223> OTHER INFORMATION: Xaa is cyclohexylalanine
205 <220> FEATURE:
206 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
207       sequence
209 <400> SEQUENCE: 7
W--> 210 Asp Glu Xaa Glu Xaa
211    1            5
214 <210> SEQ ID NO: 8
215 <211> LENGTH: 17
216 <212> TYPE: PRT
217 <213> ORGANISM: Artificial Sequence
219 <220> FEATURE:
220 <221> NAME/KEY: MOD_RES
221 <222> LOCATION: (17)
222 <223> OTHER INFORMATION: AMIDATION
224 <220> FEATURE:
225 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
226       sequence.
228 <400> SEQUENCE: 8
229 Lys Lys Lys Gly Ser Val Val Ile Val Gly Arg Ile Ile Leu Ser Gly
230    1            5              10               15
232 Arg
235 <210> SEQ ID NO: 9
236 <211> LENGTH: 13
237 <212> TYPE: PRT
238 <213> ORGANISM: Artificial Sequence
240 <220> FEATURE:
241 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
242       sequence
244 <400> SEQUENCE: 9
245 Asp Glu Met Glu Glu Cys Ala Ser His Leu Pro Tyr Lys
246    1            5                  10
249 <210> SEQ ID NO: 10
250 <211> LENGTH: 4
251 <212> TYPE: PRT
252 <213> ORGANISM: Artificial Sequence
254 <220> FEATURE:
255 <221> NAME/KEY: SITE
256 <222> LOCATION: (1)..(3)
257 <223> OTHER INFORMATION: Phenylalanines are linked by an ether bond
259 <220> FEATURE:
260 <221> NAME/KEY: SITE
261 <222> LOCATION: (4)
262 <223> OTHER INFORMATION: Xaa is 4,4-difluoro-2-amino butyric acid
264 <220> FEATURE:
265 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
266       sequence
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/719,261

DATE: 08/30/2001
TIME: 09:46:14

Input Set : A:\Mewburn-Matassa ('261) Sequence Listing.txt
Output Set: N:\CRF3\08302001\I719261.raw

```
       268 <400> SEQUENCE: 10
W--> 269 Phe Glu Phe Xaa
     -270    1
       273 <210> SEQ ID NO: 11
       274 <211> LENGTH: 6
       275 <212> TYPE: PRT
       276 <213> ORGANISM: Artificial Sequence
       278 <220> FEATURE:
       279 <221> NAME/KEY: SITE
       280 <222> LOCATION: (3)
       281 <223> OTHER INFORMATION: Xaa is diphenylalanine
       283 <220> FEATURE:
       284 <221> NAME/KEY: SITE
       285 <222> LOCATION: (5)
       286 <223> OTHER INFORMATION: Xaa is cyclohexylalanine
       288 <220> FEATURE:
       289 <221> NAME/KEY: SITE
       290 <222> LOCATION: (6)
       291 <223> OTHER INFORMATION: Xaa is 3-amino-5,5-difluoro-pentanoic acid
       293 <220> FEATURE:
       294 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
       295       sequence
       297 <400> SEQUENCE: 11
W--> 298 Asp Glu Xaa Glu Xaa Xaa
       299    1                5
       302 <210> SEQ ID NO: 12
       303 <211> LENGTH: 5
       304 <212> TYPE: PRT
       305 <213> ORGANISM: Artificial Sequence
       307 <220> FEATURE:
       308 <221> NAME/KEY: SITE
       309 <222> LOCATION: (2)
       310 <223> OTHER INFORMATION: Xaa is diphenylalanine
       312 <220> FEATURE:
       313 <221> NAME/KEY: SITE
       314 <222> LOCATION: (4)
       315 <223> OTHER INFORMATION: Xaa is cyclohexylalanine
       317 <220> FEATURE:
       318 <221> NAME/KEY: SITE
       319 <222> LOCATION: (5)
       320 <223> OTHER INFORMATION: Xaa is 4,4-difluoro-2-amino butyric acid
       322 <220> FEATURE:
       323 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
       324       sequence
       326 <400> SEQUENCE: 12
W--> 327 Glu Xaa Ile Xaa Xaa
       328    1             5
       331 <210> SEQ ID NO: 13
       332 <211> LENGTH: 6
```

**VERIFICATION SUMMARY**
PATENT APPLICATION: US/09/719,261

DATE: 08/30/2001
TIME: 09:46:15

Input Set : A:\Mewburn-Matassa ('261) Sequence Listing.txt
Output Set: N:\CRF3\08302001\I719261.raw

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:54 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:74 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:103 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:151 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7
L:269 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10
L:298 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11
L:327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:356 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13